

ON MAXIMIZING THE AVERAGE TIME AT A GOAL

S. DEMKO and T. P. HILL

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

In a decision process (gambling or dynamic programming problem) with finite state space and arbitrary decision sets (gambles or actions), there is always available a Markov strategy which uniformly (nearly) maximizes the average time spent at a goal. If the decision sets are closed, there is even a stationary strategy with the same property.

Examples are given to show that approximations by discounted or finite horizon payoffs are not useful for the general average reward problem.

gambling theory * goal problems * dynamic programming * stationary strategy
* Markov strategy * average reward criterion

1. Introduction

The subject of this paper is finite state, discrete time decision processes with single fixed goals and arbitrary decision sets (all to be defined precisely in Section 2). Various objective functions associated with such processes have been studied extensively, among them: maximizing the probability of reaching the goal [7, 8, 15, 19, 20]; minimizing the expected time or cost to the goal [4, 5, 12, 13, 16, 17]; maximizing the expected total number of times at the goal [15]; maximizing the expected total discounted rewards at the goal [2, 6, 13, 16]; maximizing the probability the goal is hit infinitely often [7, 11, 19, 20]; and maximizing the expected average time at the goal [3, 6, 10, 13, 16]. We are concerned here with this last objective; finite state goal problems with average reward criterion.

In the pioneering work of Howard [14], and much of the subsequent research, e.g. Ross [16], the assumption of finiteness of number of gambles available at each point was essential, and led to constructive determinations of an optimal stationary strategy. If the number of gambles at some states are infinite, however, optimal strategies need not exist, and stationary strategies are not at all good in general. Ross [16, p. 144] has raised the question of determining the smallest class of strategies which are ϵ -optimal for average reward problems. This problem has been studied by Chitashvili [3] and Fainberg [10]. It is the purpose of this paper to completely answer that question in the case of finite state goal problems.

It is shown in Theorem 1 that in every such problem with closed sets of gambles there always exists a stationary strategy which is uniformly ϵ -optimal; this result is used to show (Theorem 2) that in the finite state goal problem with arbitrary decision sets there always exists a Markov strategy which is uniformly nearly optimal. Examples are given to show that the relationship between these average reward problems, and problems with discounted or finite horizon objectives, is not very close.

2. Statements of results

Definition 2.1. As in [4], a finite state goal problem is a triple (X, Γ, g) , where X is a finite set, Γ associates to each point $x \in X$ a nonempty collection $\Gamma(x)$ of probability measures on X , and $g \in X$ is a distinguished element of X .

The set X represents the state space or fortune space of the process, $\Gamma(x)$ the actions or gambles available at the state x , and g the 'goal' state.

Much of the notation will follow that of Dubins and Savage [7]. The Dirac delta-measure at x will be denoted by $\delta(x)$. A strategy is a function from finite sequences in X (including the empty sequence ' \emptyset ') to probability measures on X . The same symbol, σ , will be used to denote both a strategy and the probability measure generated by σ on the product sigma-algebra on $X^{\mathbb{N}}$ (X endowed with the discrete sigma algebra). $E_{\sigma}f$ will denote the integral of f with respect to σ . A strategy σ in Γ at x , written $\sigma \in \Gamma/x$, is a strategy σ such that $\sigma(\emptyset) \in \Gamma(x)$, and $\sigma(x_1, \dots, x_n) \in \Gamma(x_n)$ for all $x_1, x_2, \dots, x_n \in X$ and all $n \in \mathbb{N}$. A strategy σ is *Markov* if $\sigma(x_1, \dots, x_n) = \sigma(x'_1, \dots, x'_n)$ whenever $x_n = x'_n$, and is *stationary* if $\sigma(x_1, \dots, x_n) = \sigma(x'_1, \dots, x'_m)$ whenever $x_n = x'_m$. The conditional strategy given x_1, \dots, x_m , $\sigma[x_1, \dots, x_m]$, is defined by $\sigma[x_1, \dots, x_m](x) = \sigma(x_1, \dots, x_m, x)$. For a function Γ from X to subsets of probability measures on X , $\bar{\Gamma}$, the *closure* of Γ , is the function defined by $\bar{\Gamma}(x) = \overline{\Gamma(x)}$ for all x , where \bar{S} is the total-variation norm closure of the set of measures S .

Definition 2.2. $N(n): X^{\mathbb{N}} \rightarrow \{0, 1, \dots, n\}$ is the function

$$N(n)(x_1, x_2, \dots) = |\{j: 1 \leq j \leq n \text{ and } x_j = g\}|,$$

that is, the number of times the goal is visited in the first n steps of the game.

Definition 2.3. The *hitting time* of $C \subset X$, T_C , is defined by $T_C(x_1, x_2, \dots) = \min\{j: x_j \in C\}$ if such a j exists, and $= \infty$ otherwise. The *time between the $n-1$ st and n th visits to the goal*, T_n , is defined inductively by $T_0 = 0$, and

$$T_n(x_1, x_2, \dots) = \min\{j: T_{n-1}(x_1, x_2, \dots) = k \text{ and } x_{k+j} = g\}$$

if such a j exists, and $= \infty$ otherwise.

Definition 2.4. For a strategy σ , the expected average time at the goal using σ , $A(\sigma)$, is defined to be $A(\sigma) = E_\sigma(\limsup N(n)/n)$, and the maximal expected average time at the goal starting at state x , $A(x)$, is $A(x) = \sup\{A(\sigma): \sigma \in \Gamma/x\}$.

Definition 2.5. Let A_1, A_2, A_3 be the functions

$$A_1(\sigma) = E_\sigma \liminf_n (N(n)/n), \quad A_2(\sigma) = \liminf_n E_\sigma(N(n)/n) \quad \text{and} \\ A_3(\sigma) = \limsup_n E_\sigma(N(n)/n),$$

and similarly let

$$A_i(x) = \sup\{A_i(\sigma): \sigma \in \Gamma/x\}.$$

Remark. In most of the previous research on average reward problems the payoffs A_2 and A_3 have been used, apparently for mathematical expediency. The average reward criterion $A(\sigma)$ defined above seems more natural to the authors, and in any event it turns out (Corollary following Theorem 2) that these criteria are equivalent for finite state problems.

The following theorem, the main result of this paper, states that in every finite state goal problem with closed (hence compact) sets of gambles, there always exists a stationary strategy which is uniformly nearly optimal.

Theorem 1. If (X, Γ, g) is any goal problem with $|X| < \infty$ and $\Gamma(x)$ closed for all x , then for each $\varepsilon > 0$ there is a stationary strategy σ^ε in Γ satisfying $A(\sigma^\varepsilon[x]) \geq A(x) - \varepsilon$ for all $x \in X$.

A version of Theorem 1 for the payoff A_2 is a special case of a result of Chitashvili [3] and one for payoff A_3 is a special case of the work of Fainberg [10]. (See also Chapter 7 of Dynkin and Yushkevich [9].) These authors consider the general average reward problem for finite X and compact $\Gamma(x)$ and exploit the relationship between the β -discounted payoff and the average reward payoff. This approach won't work for payoff A . As the following example, which is similar to example 3 of Bather [2], shows even if the limit of good (discounted or finite horizon) strategies exist, this limiting strategy may be worthless in all respects.

Example. $X = \{a, b, g\}$;

$$\Gamma(g) = \{\delta(g)\}, \quad \Gamma(b) = \{\delta(b)\}, \\ \Gamma(a) = \{\delta(g)/n + \delta(b)/n^2 + (1 - 1/n - 1/n^2)\delta(a): n \geq 1\} \cup \{\delta(a)\}.$$

As $\beta \rightarrow 1$, the limit of good strategies for the β -discounted reward problem exists, and in fact is the stationary strategy which uses $\delta(a)$ at state 'a'. Similarly for the

limit, as $n \rightarrow \infty$, of strategies which are good for the n -step problem. But using $\delta(a)$ always at state 'a' will *never* lead to the goal.

Theorem 2. *If (X, Γ, g) is any goal problem with $|X| < \infty$, then for each $\varepsilon > 0$ there is a Markov strategy σ^m in Γ satisfying $A(\sigma^m[x]) \geq A(x) - \varepsilon$ for all $x \in X$.*

Theorem 1 and the proof of Theorem 2 yield the following corollary.

Corollary. *If (X, Γ, g) is a goal problem with $|X| < \infty$, then*

$$A(x) = A_1(x) = A_2(x) = A_3(x) \quad \text{for all } x \in X.$$

3. Proofs

The proof of Theorem 1 requires several lemmas. The first is an easy exercise.

Lemma 1

$$\limsup_{m \rightarrow \infty} m / (T_1 + \dots + T_m) = \limsup_{n \rightarrow \infty} \frac{N(n)}{n}.$$

Lemma 2. *If $|X| < \infty$ and $\Gamma = \tilde{\Gamma}$, then*

$$\lim_{M \rightarrow \infty} [\inf\{E_\sigma(T_1 \wedge M) : \sigma \in \Gamma/x\}] = \inf\{E_\sigma(T_1) : \sigma \in \Gamma/x\};$$

here, $a \wedge b := \min(a, b)$.

Proof. ' \leq ' Trivial, since $E_\sigma(T_1 \wedge M) \leq E_\sigma(T_1)$ for all σ and all M .

' \geq ' Let $\lim_{M \rightarrow \infty} [\inf\{E_\sigma(T_1 \wedge M) : \sigma \in \Gamma/x\}] = a$. If $a = +\infty$, then we are done, so suppose $a < \infty$ and fix $\varepsilon > 0$. Note that

$$\inf\{E_\sigma(T_1 \wedge n) : \sigma \in \Gamma/x\} \text{ is nondecreasing in } n, \text{ and bounded above by } a. \quad (1)$$

For each $n = 1, 2, \dots$ pick $\sigma_n \in \Gamma/x$ satisfying

$$E_{\sigma_n}(T_1 \wedge n) < a + \varepsilon. \quad (2)$$

Since the number of finite sequences in X is countable, and $\Gamma(x)$ is compact for all x , there exists a subsequence $\{\sigma_{n_k}\}$ and $\sigma \in \Gamma/x$ such that

$$\lim_{k \rightarrow \infty} \sigma_{n_k} = \sigma. \quad (3)$$

Since ε was arbitrary, the proof will be complete once it is shown that

$$E_\sigma(T_1) < a + 3\varepsilon. \quad (4)$$

Case 1. $E_\sigma(T_1) < \infty$. Pick M so that

$$E_\sigma(T_1 \wedge M) \geq E_\sigma(T_1) - \varepsilon. \quad (5)$$

By (3) we can pick $K \geq M$ so that

$$E_{\sigma_K}(T_1 \wedge M) \geq E_\sigma(T_1 \wedge M) - \varepsilon. \quad (6)$$

Since $K \geq M$, it follows that

$$E_{\sigma_K}(T_1 \wedge K) \geq E_{\sigma_K}(T_1 \wedge M). \quad (7)$$

Hence

$$a + \varepsilon > E_{\sigma_K}(T_1 \wedge K) \geq E_{\sigma_K}(T_1 \wedge M) \geq E_\sigma(T_1 \wedge M) - \varepsilon \geq E_\sigma(T_1) - 2\varepsilon,$$

where the inequalities follow from (2), (7), (6) and (5) respectively. This completes Case 1.

Case 2. $E_\sigma(T_1) = \infty$. Pick M so that $E_\sigma(T_1 \wedge M) \geq a + 1$, and pick $K \geq M$ as in (6). Then

$$a + \varepsilon \geq E_{\sigma_K}(T_1 \wedge K) \geq E_{\sigma_K}(T_1 \wedge M) - \varepsilon \geq E_\sigma(T_1 \wedge M) - \varepsilon a + 1 - \varepsilon$$

(by (2), (7), (6) and choice of M), a contradiction for $\varepsilon < \frac{1}{2}$. Thus $E_\sigma(T_1) < \infty$ (assuming $a < \infty$) and case 1 applies. This completes the proof. \square

Lemma 3. Let M be finite and define $T'_i := T_i \wedge M$. If $E_\sigma(T'_i) \geq a$ for all σ in Γ at g , then

$$\liminf_{n \rightarrow \infty} (T'_1 + \cdots + T'_n)/n \geq a \quad \text{a.s. } \sigma \text{ for all } \sigma \text{ in } \Gamma \text{ at } g.$$

Proof. Fix σ in Γ at g . Let $F_0 = \{\phi, X^\mathbb{N}\}$, and for $n \geq 1$, let F_n be the sigma-field generated by T'_1, \dots, T'_n . Then $Y_i = T'_i - E_\sigma(T'_i | F_{i-1})$ is a martingale difference sequence satisfying $|Y_i| \leq T'_i \leq M$ almost surely. Consequently, $(1/n) \sum_{i=1}^n Y_i \rightarrow 0$ almost surely. Since $E_\sigma(T'_i | F_{i-1}) \geq a$ almost surely, it follows that

$$\liminf_{n \rightarrow \infty} (T'_1 + \cdots + T'_n)/n = \liminf_{n \rightarrow \infty} \{(E_\sigma(T'_1 | F_0) + \cdots + E_\sigma(T'_n | F_{n-1}))/n\} \geq a \quad \text{a.s. } \sigma. \quad \square$$

Lemma 4. If $|X| < \infty$ and $\Gamma = \bar{\Gamma}$, then

$$\limsup_{n \rightarrow \infty} N(n)/n \leq [\inf\{E_\sigma(T_1) : \sigma \in \Gamma/g\}]^{-1} \quad \text{a.s. } \sigma \text{ for all } \sigma \text{ in } \Gamma \text{ at } g.$$

Proof. Let $\inf\{E_\sigma(T_1) : \sigma \in \Gamma/g\} = a$.

Case 1 $a < \infty$. Fix $\varepsilon > 0$. By Lemma 2 it is possible to find $M \in \mathbb{N}$ so that $\inf\{E_\sigma(T_1 \wedge M) : \sigma \in \Gamma/g\} \geq a - \varepsilon$. Now define $T'_i := T_i \wedge M$. We have, by Lemma 3,

$$\liminf_{n \rightarrow \infty} (T_1 + \cdots + T_n)/n \geq \liminf_{n \rightarrow \infty} (T'_1 + \cdots + T'_n)/n \geq a - \varepsilon$$

almost surely for all σ in Γ at g . Thus, by Lemma 1,

$$\limsup_{n \rightarrow \infty} \frac{N(n)}{n} \leq (a - \varepsilon)^{-1} \quad (8)$$

almost surely for all σ in Γ at g . Since $\varepsilon > 0$ was arbitrary, the desired conclusion follows from (8).

Case 2 $a = \infty$. Then $E_\sigma(T_1) = \infty$ for all $\sigma \in \Gamma/g$, and it follows easily from Lemmas 1, 2 and 3 as in Case 1 that $\lim N(n)/n = 0$ a.s. for all $\sigma \in \Gamma/g$. \square

Proof of Theorem 1. Let $\inf\{E_\sigma(T_1) : \sigma \in \Gamma/g\} = a$. If $a = +\infty$, then by Lemma 4 $A(g) = 0$ and hence $A(x) = 0$ for all $x \in X$. But then every strategy, in particular every stationary strategy, is (trivially) optimal.

Suppose $a < \infty$ and fix $\varepsilon > 0$. By [4, Theorem 4.1] there is a stationary strategy $\sigma^{\alpha'}$ in Γ satisfying

$$E_{\sigma^{\alpha'}|g}(T_1) \leq a + \varepsilon. \quad (9)$$

Let $C \subset X$ be the closed communicating class relative $\sigma^{\alpha'}$ containing g (which is clearly recurrent). By [7, Theorem 3.9.2] there is a stationary strategy $\sigma^{\alpha''}$ in Γ satisfying

$$\sigma^{\alpha''}[x](T_C < \infty) \geq \sup\{\sigma(T_C < \infty) : \sigma \in \Gamma/x\} - \varepsilon \quad \text{for all } x \in X. \quad (10)$$

Define the stationary strategy σ^x in Γ by $\sigma^x(x) = \sigma^{\alpha'}(x)$ if $x \in C$, and $= \sigma^{\alpha''}(x)$ if $x \notin C$. The remainder of the proof consists in showing that for any strategy σ in Γ at x , σ^x satisfies

$$A(\sigma^x[x]) \geq A(\sigma) - 2\varepsilon. \quad (11)$$

Fix $\sigma \in \Gamma/x$. By [16, Theorem 4.7], (9), (10) and the fact that $a \geq 1$, it follows that

$$\begin{aligned} A(\sigma^x[x]) &= \sigma^x[x](T_C < \infty)[E_{\sigma^{\alpha'}|g}(T_1)]^{-1} \geq \sigma^{\alpha''}[x](T_C < \infty) \cdot a^{-1} - \varepsilon \\ &\geq \sigma(T_C < \infty) \cdot a^{-1} - 2\varepsilon. \end{aligned}$$

Since $A(\sigma) \leq \sigma(T_1 < \infty) \cdot A(g) \leq \sigma(T_C < \infty) \cdot A(g)$, and since $0 \leq A \leq 1$, the inequality in (11) follows by Lemma 4. This completes the proof. \square

In Howard's treatise [14], where $I'(x)$ is finite for all x , a 'policy-improvement' algorithm is given which converges to that stationary strategy which is optimal among all stationary strategies. Theorem 1 implies that the resulting stationary strategy of Howard's technique is optimal even among *all* strategies.

For the proof of Theorem 2 we use

Lemma 5. *If $|X| < \infty$, then for every $\bar{\sigma} \in \bar{\Gamma}/x$ and each $\varepsilon > 0$ there is a $\sigma \in \Gamma/x$ with $\|\sigma - \bar{\sigma}\| < \varepsilon$. Moreover, if $\bar{\sigma}$ is stationary, then σ can be chosen to be Markov.*

Proof. Fix $\varepsilon > 0$, and pick $\sigma(\emptyset)$ so that $\|\sigma(\emptyset) - \bar{\sigma}(\emptyset)\| < \varepsilon/2$. For each finite sequence x_1, x_2, \dots, x_n of elements in X , pick $\sigma(x_1, \dots, x_n) \in \Gamma(x_n)$ so that $\|\sigma(x_1, \dots, x_n) - \bar{\sigma}(x_1, \dots, x_n)\| < \varepsilon/2^{n+1}|X|^{n+1}$. If $\bar{\sigma}$ is stationary, pick $\sigma(x_1, \dots, x_n)$ so that $\sigma(x_1, \dots, x_n) = \sigma(x'_1, \dots, x'_n)$ whenever $x_n = x'_n$. It is easy to verify that the strategy σ so chosen satisfies $\|\bar{\sigma} - \sigma\| < \varepsilon$, and is Markov if $\bar{\sigma}$ is stationary. \square

Proof of Theorem 2 completed. Fix $\varepsilon > 0$, and let \bar{A} be that for the problem $(X, \bar{\Gamma}, g)$, that is, $\bar{A}(x) = \sup\{A(\sigma) : \sigma \in \bar{\Gamma}/x\}$. By Theorem 2 there is a stationary strategy σ^∞ in $\bar{\Gamma}$ satisfying $A(\sigma^\infty[x]) \geq \bar{A}(x) - \varepsilon$ for all $x \in X$. By Lemma 5 there exists a Markov strategy σ^m in Γ with $\|\sigma^m - \sigma^\infty\| < \varepsilon$. Since $0 \leq A \leq \bar{A} \leq 1$, σ^m satisfies

$$A(\sigma^m[x]) \geq A(\sigma^\infty[x]) - \varepsilon \geq \bar{A}(x) - 2\varepsilon \geq A(x) - 2\varepsilon$$

for all $x \in X$. \square

Proof of the Corollary. Observe that, by Fatou's Lemma, $A_1 \leq A_2 \leq A_3 \leq A$. Next, as in the proof of Theorem 2, we can find a stationary strategy σ^∞ in $\bar{\Gamma}$ which is uniformly (nearly) optimal for \bar{A} . But for stationary strategies, $\lim N(n)/n$ exists almost surely, and application of the Dominated Convergence Theorem shows $\bar{A} = \bar{A}_1 = \bar{A}_2 = \bar{A}_3$. Lemma 5 implies that $A = \bar{A}$, and $A_i = \bar{A}_i$ for $i = 1, 2, 3$, and this completes the proof. \square

Remarks. Although $A(x) = A_1(x) = A_2(x) = A_3(x)$ for all x if $|X| < \infty$, it is easy to see that $|X| < \infty$ does not imply $A(\sigma) = A_1(\sigma) = A_2(\sigma) = A_3(\sigma)$ for all σ in Γ . However, the proof of Theorem 2 shows that there is always a Markov strategy which is uniformly (nearly) optimal with respect to A, A_1, A_2 , and A_3 simultaneously. In fact the ε -optimal strategies constructed in Theorems 1 and 2 are even easily seen to be *persistently ε -optimal* [8], since uniformly good stationary strategies are automatically persistently good.

4. Other reward criteria

This paper leaves open the question of whether stationary strategies (in the $|X| < \infty, \Gamma = \bar{\Gamma}$ case) and Markov strategies (in the general $|X| < \infty$ case) are uniformly adequate for the more general average reward problem in which reward $r(x)$ is obtained at each visit to state x . They suspect the answer is affirmative, but the techniques used in these proofs do not even carry over in the special case of a goal set $G \subset X$ (i.e. $r(x) = 1$ if $x \in G, = 0$ otherwise), even if $\Gamma = \bar{\Gamma}$.

Example. $X = \{a, b, g, g'\}$;

$$\Gamma(a) = \{(\delta(b) + \delta(g))/2\} \cup \{\delta(g')\}, \quad \Gamma(b) = \{\delta(b)\},$$

$$\Gamma(g) = \{\delta(g)\}, \quad \Gamma(g') = \{\delta(b)\};$$

and $G = \{g, g'\}$. If one is at state a , and wishes to minimize his expected time to G , or maximize his probability of hitting G (approaches used in the above proofs), he uses $\delta(g')$, which is bad for maximizing the average time in G .

The relationship between the average reward criterion, and discounted or finite horizon reward payoffs seems to be rather weak, perhaps because for the average reward problem, a gambler is not penalized for 'resting' at neutral states for long periods initially, as long as he eventually makes good decisions later. In discounted, or finite horizon problems, on the other hand, the gambler is penalized heavily for staying at worthless states for long periods initially, but is not penalized much for making bad decisions in the distant future. Neither the ε -optimal stationary strategies guaranteed by Blackwell's results [2] for the discounted reward problem, nor the ε -optimal Markov strategies found by backward induction [6, 13] seemed useful in analyzing the average reward problem.

Acknowledgement

The authors would like to thank Professors L. Dubins, S. Ross, and W. Sudderth for several useful conversations, especially Sudderth who first mentioned this problem to Hill in a seminar in Berkeley in 1976.

References

- [1] J. Bather, Optimal decision procedures for finite Markov chains, *Adv. Appl. Prob.* 5 (1973) 328–339.
- [2] D. Blackwell, Discounted dynamic programming, *Ann. Math. Statist.* 36 (1965) 226–235.
- [3] R. Chitashvili, A controlled finite Markov chain with an arbitrary set of decisions, *Theory of Probability and Its Applications* 20 (1975) 839–846.
- [4] S. Demko and T. Hill, Decision processes with total-cost criteria, *Ann. Prob.* 9 (1981) 293–301.
- [5] C. Derman, On sequential decisions and Markov chains, *Management Sci.* 9 (1962) 16–24.
- [6] C. Derman, *Finite State Markovian Decision Process* (Academic Press, New York, 1970).
- [7] L. Dubins and L. Savage, *Inequalities for Stochastic Processes* (Dover, New York, 1976).
- [8] L. Dubins and W. Sudderth, Persistently ε -optimal strategies, *Math. Operations Research* 2 (1977) 125–135.
- [9] E. Dynkin and A. Yushkevich, *Controlled Markov Processes* (Springer-Verlag, Berlin, Heidelberg, New York, 1979).
- [10] E. Fainberg, The existence of a stationary π -optimal policy for a finite Markov chain, *Theory of Probability and Its Applications* 23 (1978) 297–313.
- [11] T. Hill, On the existence of good Markov strategies, *Trans. Amer. Math. Society* 247 (1979) 157–176.
- [12] T. Hill, On reaching a goal quickly, AFOSR Technical Report, Georgia Institute of Technology, 1979.
- [13] A. Hordijk, *Dynamic Programming and Markov Potential Theory*, Math. Centre Tracts, Amsterdam, 1974.
- [14] R. Howard, *Dynamic Programming and Markov Processes* (Technology Press, Cambridge, Mass., 1960).

- [15] D. Ornstein, On the Existence of stationary optimal strategies, Proc. Amer. Math. Society 20 (1969) 563–569.
- [16] S. Ross, Applied Probability Models with Optimization Applications (Holden-Day, San Francisco, 1970).
- [17] M. Schäl, On dynamic programming. Compactness of the space of policies, Stochastic Proc. Appl. 3 (1975) 345–364.
- [18] W. Stout, Almost Sure Convergence (Academic Press, New York, 1974).
- [19] W. Sudderth, On the existence of good stationary strategies, Trans. Amer. Math. Society 135 (1969) 399–414.
- [20] W. Sudderth, On measurable gambling problems, Ann. Math. Statist. 42 (1969) 260–269.